Neuroscience Research 78 (2014) 90–94

Contents lists available at ScienceDirect

Neuroscience Research

journal homepage: www.elsevier.com/locate/neures

Rapid communication

Saliency-guided neural prosthesis for visual attention: Design and simulation

Masatoshi Yoshida[a,b,*], Richard Veale[c]

[a] Department of Developmental Physiology, National Institute for Physiological Sciences, Okazaki, Japan
[b] School of Life Science, The Graduate University for Advanced Studies, Hayama, Japan
[c] Cognitive Science Program, Indiana University, Bloomington, USA

ARTICLE INFO

Article history:
Received 29 March 2013
Received in revised form 8 July 2013
Accepted 18 July 2013
Available online 3 August 2013

Keywords:
Computational neuroscience
Superior colliculus
Electrical microstimulation

ABSTRACT

Recently the authors showed that a computational model of visual saliency could account for changes in gaze behavior of monkeys with damage in the primary visual cortex. Here we propose a neural prosthesis to restore eye gaze behavior by electrically stimulating the superior colliculus to drive visual attention. The saliency computational model is used to calculate the stimulation parameters from a real-time camera stream. Our simulations demonstrate that electrodes implanted in the superior colliculus at 1.0 mm spacing are, in principle, able to recover specifically those visual attention behaviors which are lost when the primary visual cortex is damaged.

There are several prosthetic options available to blind persons for the recovery of their functional vision. For example, one approach evokes visual experiences ("phosphenes") via electrical stimulation of primary visual cortex (V1) (Dobelle and Mladejovsky, 1974; Schmidt et al., 1996; Bradley, 2004). However, this approach is not applicable to patients with damage in V1 and has several disadvantages, including the difficulty of implanting electrodes precisely in visual cortex, which is large and folded. In this paper, we propose a different approach in which electrical stimulation in the superior colliculus is used to drive attention to saliency rather than directly evoke visual experiences.

A computational model cal

2001) quantifies the influence o
tion. The model has been used to predict human eye movements (Itti, 2005) and to diagnose neurological patients based on abnormal eye movement patterns (Tseng et al., 2012). Recently the authors showed that the model also predicts gaze in V1-damaged monkeys during free-viewing (Yoshida et al., 2012). We examined which visual features were most important to guide eye movements in monkeys with damage to the visual cortex. We found that motion, luminance and color features contributed to residual guidance of eye movement, whereas orientation feature did not. This directly motivates the current research: *is it possible to artificially restore the influence of features such as orientation on gaze?*

We propose a saliency-based visual prosthesis in which a brain region is electrically stimulated to recover gaze behavior in visually impaired patients. The superior colliculus (SC) and frontal eye field (FEF) are candidates for electrode implantation because (1) both areas are implicated in saliency computation (Navalpakkam and Itti, 2005; Thompson and Bichot, 2005; Fecteau and Munoz, 2006) and (2) electrical microstimulation to these areas affects covert attention without evoking eye movements (Moore and Fallah, 2001; Moore, 2003; Carello and Krauzlis, 2004; Müller et al., 2005)

sents visual space in a 4 mm by 3 mm area, making it easier to cover the whole structure with electrodes. Furthermore, the retinotopic map and spatial interactions with SC are well established.

We now describe the overall design of the prosthetic system as we expect to implement it in future physiology experiments. The proposed prosthesis operates in a tight loop with five steps (Fig. 1). A camera captures the field of view of the subject using eye-tracking, and a saliency map for that image is computed using a saliency model encoding the visual attention functions that we wish to recover (steps 1–3). This saliency map is used to compute optimal stimulation parameters for electrodes to maximally recover the saliency map's pattern of activity in SC (steps 4 and 5). Each step is detailed below.

Step 1: A field-of-view (FOV) image is acquired via eye tracking.

Step 2: The saliency map of the FOV image is computed. These computations can be accomplished in as little as 2 ms per frame (Chang et al., 2010), sufficiently fast for a real-time system.

* Corresponding author at: Department of Developmental Physiology, National Institute for Physiological Sciences, Okazaki, Japan. Tel.: +81 564557764.
E-mail address: myoshi@nips.ac.jp (M. Yoshida).

0168-0102
http://dx.doi.org/10.1016/j.neures.2013.07.007

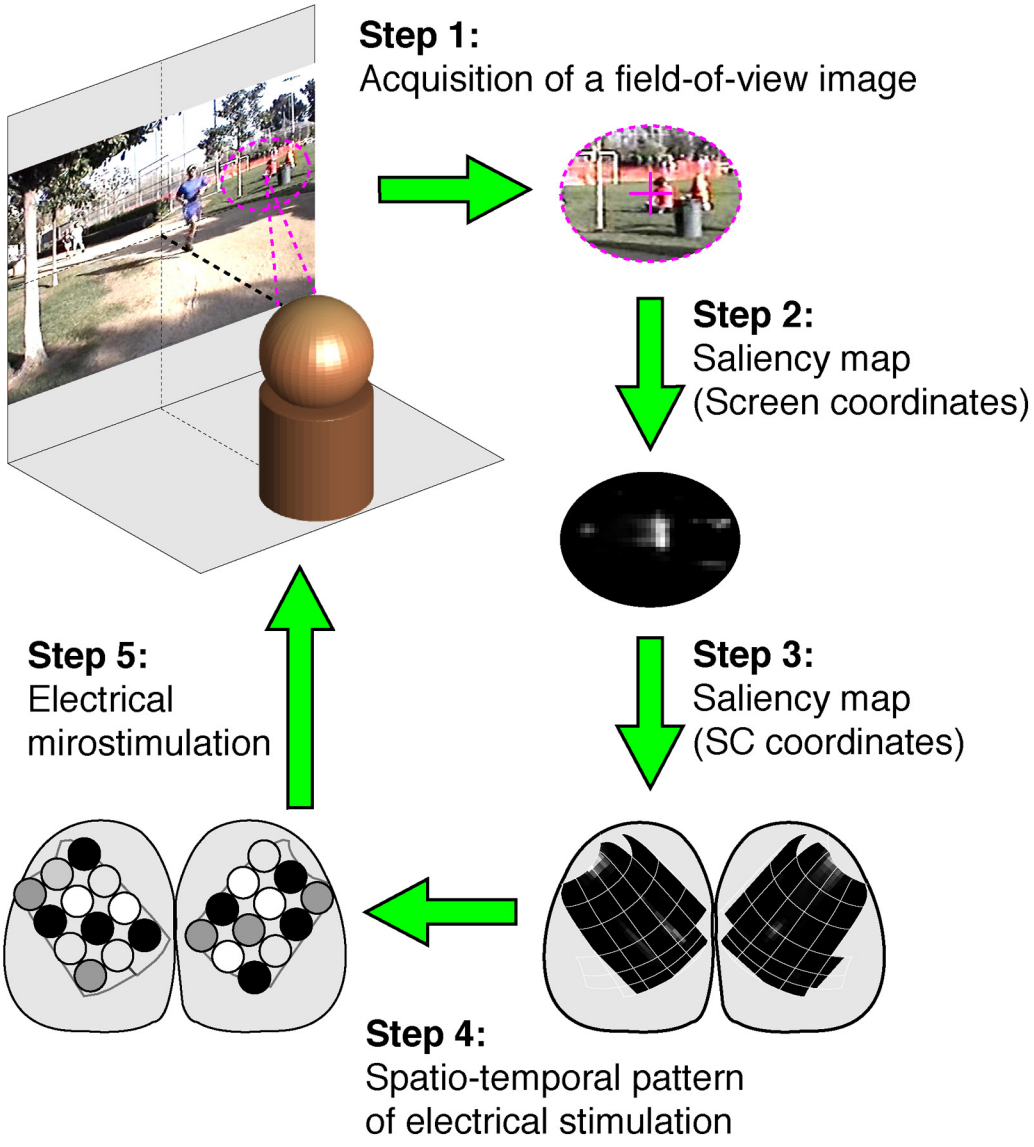

Fig. 1. Overview of the system.

Step 3: The FOV saliency map is transformed to the superior colliculus coordinates using known relationships between anatomical SC location and points in the visual field (Ottes et al., 1986).

Step 4: Using the SC-space saliency map, the spatio-temporal pattern of electrical stimulation that will maximally reproduce desired activity is computed. This is accomplished via a model which simulates the response of SC.

Step 5: The electrical stimulation is applied via electrodes implanted in SC. Since the SC is located in the depth of the brain, it is difficult to implant a multielectrode array (e.g., Utah array). We will implant penetrating electrodes that are independently controlled by microdrives (e.g., Flex MT by Alpha Omega or the Deep Vertical Microdrive by Gray Matter Research). The minimal spacing of electrodes by the microdrives ranges 0.8–1.5 mm. The intensity and frequency of stimulation currents will be adjusted so that they do not directly induce eye movements. Once the optimal current intensity is determined, we will modulate only the frequency.

We performed simulation studies to determine the optimal spacing of electrodes to reconstruct saliency maps in SC coordinates (Figs. 1 and 2C). To do this, we simulated the effect of the prosthesis on a simple neural field model based on Trappenberg's model (Trappenberg et al., 2001), which is described in Step 4 below. We now present the model used to simulate the effect of the prosthesis on the superior colliculus, and thus approximate gaze behavior. Steps 1–5 are analogs to Steps 1–5 describing the real prosthesis proposed above.

Step 1 and 2: To represent the FOV images, we randomly selected 1000 frames from videos used in monkey experiments (Yoshida et al., 2012), and computed their orientation channel saliency maps.

Step 3: The value of each SC-space pixel is the integral of the values of parts of field-of-view-space pixels which fall inside a circular receptive field. The receptive field centers are computed using equations in Ottes et al. (1986) and converted to image space. The size of the screen and the distance from the screen to eyes are the same as those from previous monkey experiments with the same videos (Yoshida et al., 2012). The receptive field size as a function of eccentricity is approximated based on empirical data from a previous paper (Wallace et al., 1997). The radius r of the receptive field was expressed as $r = 0.2R + 0.5$, where R was the eccentricity of the receptive field center (in degrees).

Step 4: We implemented a stable-state approximation of Trappenberg's dynamic SC model. The model is used to compute the optimal electrode currents expected to reconstruct a pattern of activity in SC (in this step). The same model is also used to

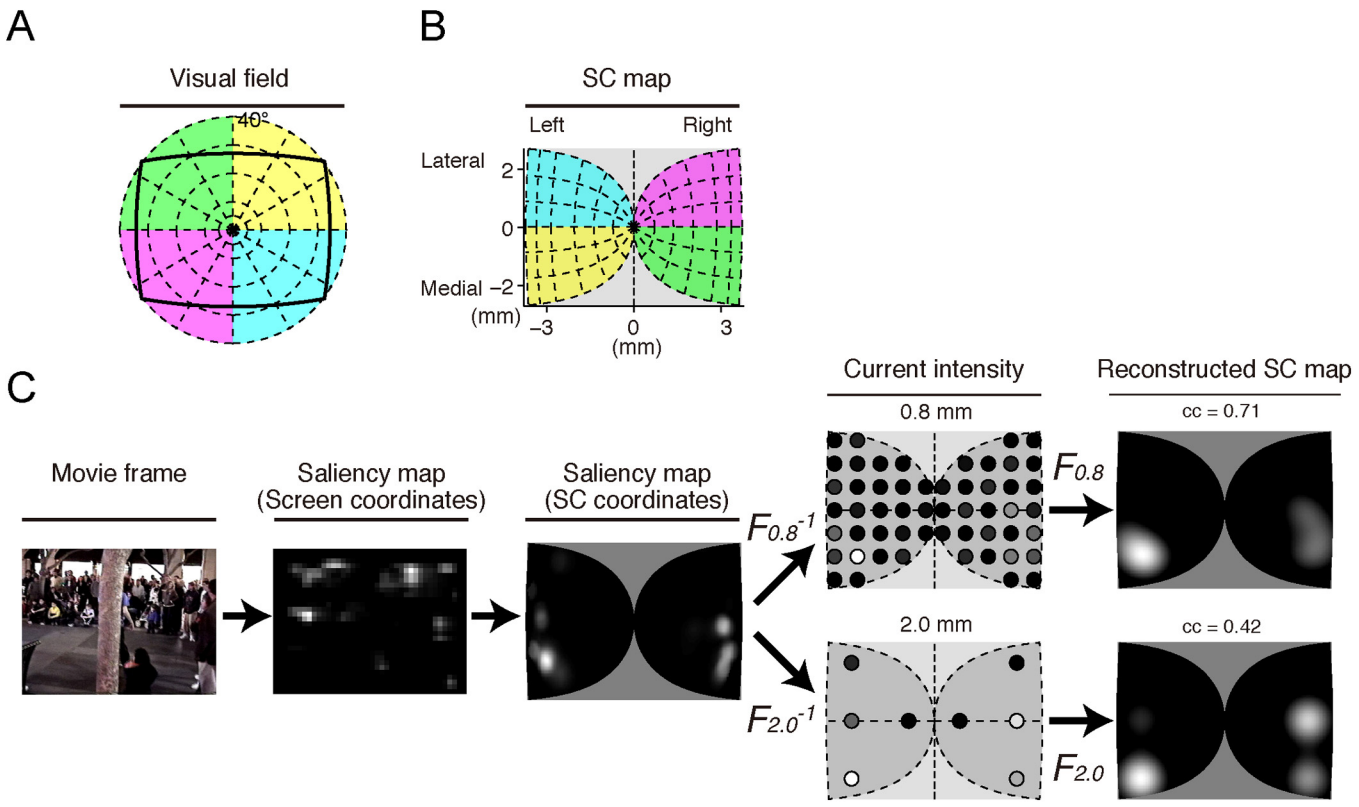

Fig. 2. Simulation examples. (A) The visual field with the position of display screen (black solid lines). (B) The SC map according to Ottes et al. (1986). The colors denote the corresponding visual field from (A). (C) An example of a video frame, its saliency map in the screen coordinates and in SC coordinates, the current intensities for two electrode spacings and the reconstructed SC map evoked by the electrical stimulation. Brighter pixels denote higher saliency or higher SC activity. Brighter circles denote higher current intensity. *F* denotes the matrix that transforms the current intensity to the SC map. The subscript number denotes the electrode spacing in millimeters. The number above the reconstructed SC map denotes the pixel-by-pixel correlation coefficients ('cc') between the saliency map in SC coordinates and the reconstructed SC maps.

simulate the actual result of that pattern of electrode stimulation (in Step 5, next). SC is represented as a matrix of very small square "neural fields" which represent the aggregate neural activity within that region (medial-lateral width 5.4 mm; anterior–posterior width 3.8 mm in 0.05 mm bins for both directions – see Fig. 2B). The activation of each neural field is computed as the sum of the influences from each of the electrodes implanted in SC. Following Trappenberg, the portion of a neural field's activation that is contributed by a given electrode is modeled as the difference of three independent Gaussian functions centered on the electrode. These Gaussian functions represent local excitation, local inhibition, and global inhibition. To account for the dynamics in Trappenberg's model, we fit parameters of these Gaussians to best approximate the average interaction strength plot in Trappenberg et al. (2001). Formally, in our model the activation *A* in a neural field as a result of electrical stimulation by an electrode distance *D* mm away is:

$$A = a\exp\left(\frac{-D^2}{2\sigma_a^2}\right) + b\exp\left(\frac{-D^2}{2\sigma_b^2}\right) + c$$

where $a = 1.35$, $b = -1/5$, $c = -1/9$, $\sigma_a = 0.6$, $\sigma_b = 1.8$. Evoked activation in SC neural fields is always positive-rectified.

To calculate the current intensity for the electrodes we precompute a matrix *F* which converts current intensity to SC response based on our version of Trappenberg's model. To construct the matrix *F*, one electrode's current intensity is set to 1.0, all others are set to zero, and the resulting SC activity is computed. This is repeated for each electrode, resulting in one 2D matrix for each electrode we have implanted. This matrix represents the response of each of the SC's neural fields to that electrode's unit stimulation. These matrices are reshaped into $1 \times n$ arrays, and then these arrays are concatenated to form a new 2D matrix (*F*). *F*'s rows represent electrodes and *F*'s columns represent the response of a given neural field in the SC model to that electrode's unit stimulation (see equation below). In *F*, x_{ij} denotes the SC field response at position j $(1-n)$ to the electrical stimulation applied by the electrode i $(1-m)$. Note that each configuration of implanted electrodes will have a different matrix *F* (for example, $F_{0.8}$ and $F_{2.0}$ in Fig. 2C).

$$\begin{matrix}\text{Current intensity}\\ \begin{matrix}\text{electrode 1}\\ \text{electrode 2}\\ \cdots\\ \text{electrode m}\end{matrix}\end{matrix}\begin{pmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \cdots & & & \\ 0 & 0 & \cdots & 1\end{pmatrix} * F$$

$$\begin{matrix}\text{SC response}\\ =\end{matrix}\begin{pmatrix}x_{11} & x_{12} & \cdots & x_{1n}\\ x_{21} & x_{22} & \cdots & x_{2n}\\ \cdots & & & \\ x_{m1} & x_{m2} & \cdots & x_{mn}\end{pmatrix} \quad (= F)$$

Given the matrix *F*, it is then possible to compute the optimal pattern of current intensity *E* to apply to the electrodes to reconstruct a saliency map in SC coordinates ('*SC saliency map*'). The current intensity *E* is computed as $E = SC\ saliency\ map \times F^{-1}$, where F^{-1} is the pseudoinverse of *F*. Then, negative currents in *E* are positive-rectified.

Step 5: Finally, we simulate the actual SC response to the computed electrode stimulation pattern *E* from Step 4, so that we can compare it to the original saliency map and measure how well we have reconstructed it. The reconstructed SC map (Fig. 2C) is computed as $SC\ saliency\ map = E \times F$, where *F* is as described above.

Using the methodology outlined in Steps 1–5 above, we ran simulations over a large sampling of different electrode spacings to determine the optimal spacing to reconstruct saliency maps. Electrodes were placed in a square grid centered on each side of the SC. We simulated inter-electrode spacing varying from 6 mm to 0.5 mm. Fig. 2C shows an example of the visual stimuli (Fig. 2C, 'Movie frame') and the saliency map for the stimuli (Fig. 2C, 'Saliency map (Screen coordinates)') that consider only the orientation feature. The saliency map is transformed into SC coordinates (Fig. 2C, 'Saliency map (SC coordinate)'), assuming that the gaze is always on the center of the screen. Furthermore, it shows the calculated current intensity for two different electrode spacings (Fig. 2C, 'Current intensity') and the corresponding reconstructed SC maps (Fig. 2C, 'Reconstructed SC map'). To evaluate the performance of reconstruction, we computed the pixel-wise correlation coefficients between the saliency map in SC coordinates (Fig. 2C, 'Saliency map (SC coordinates)') and the reconstructed SC map (Fig. 2C, 'Reconstructed SC map') for each of 1000 frames randomly drawn from the videos used in previous experiments. The results for all electrode spacings are plotted in Fig. 3 (solid lines). The results suggest that electrodes implanted in SC with 1.0 mm spacing, which is practically possible for future implementation, are able to reconstruct major aspects of real video frame stimuli (the median of the pixel-wise correlation coefficients = 0.64). A downturn in performance at small spacings (Fig. 3, solid lines) hinted that the rectification of negative currents in Step 4 was negatively influencing the results. To verify this, we ran experiments in which the current intensity was not positive-rectified (Fig. 3, dashed lines). As expected, reintroduction of negative currents allowed the recovery of performance for small electrode spacings. This is a limitation of the linear kernel approach, which was chosen here for computational efficiency. In the future it may be necessary to use non-linear or machine learning methods to better compensate for such constraints. Finally, to demonstrate that the location of salient pixels does not affect our ability to reconstruct a saliency map, we classified the 1000 test frames into two groups based on the location of the majority of the salient pixels (edge of the frame versus the interior). The performance of reconstruction for the two groups is overall similar, suggesting that our results are not an artifact of a skewed spatial distribution of salience in our test images.

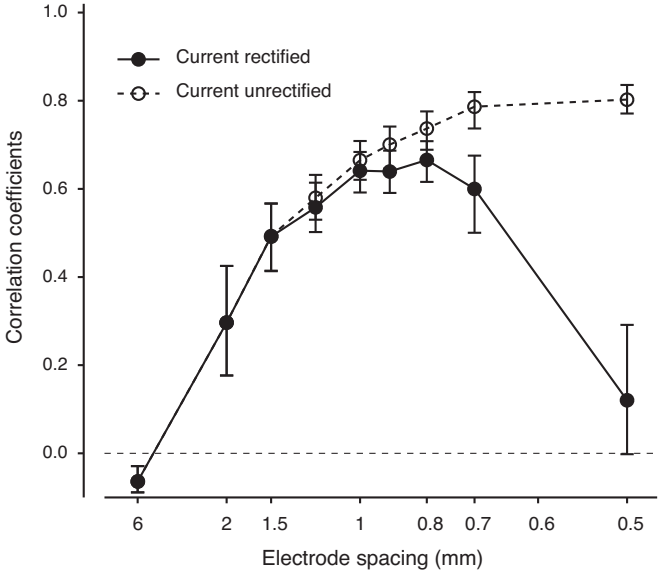

Fig. 3. Summary of the simulation results. The pixel-wise correlation coefficient between the evoked SC activity and the SC-space saliency map is plotted across different spacings of the electrodes (from 6 mm to 0.5 mm). The median values for 1000 data points are plotted. Error bars denote the 25th and 75th percentiles. The solid lines indicate simulations where negative currents are rectified and dotted lines for those in which negative currents are not rectified. Note that a reciprocal scale is used for the horizontal axis.

The simulation studies above demonstrate that the neural prosthesis is in principle able to provide information about salient positions in the visual field by modulating the SC activity pattern. Specifically, we have demonstrated that it is able to provide salience information relating to the orientation feature. Our previous study demonstrated that gaze behavior of V1-lesioned monkeys could be

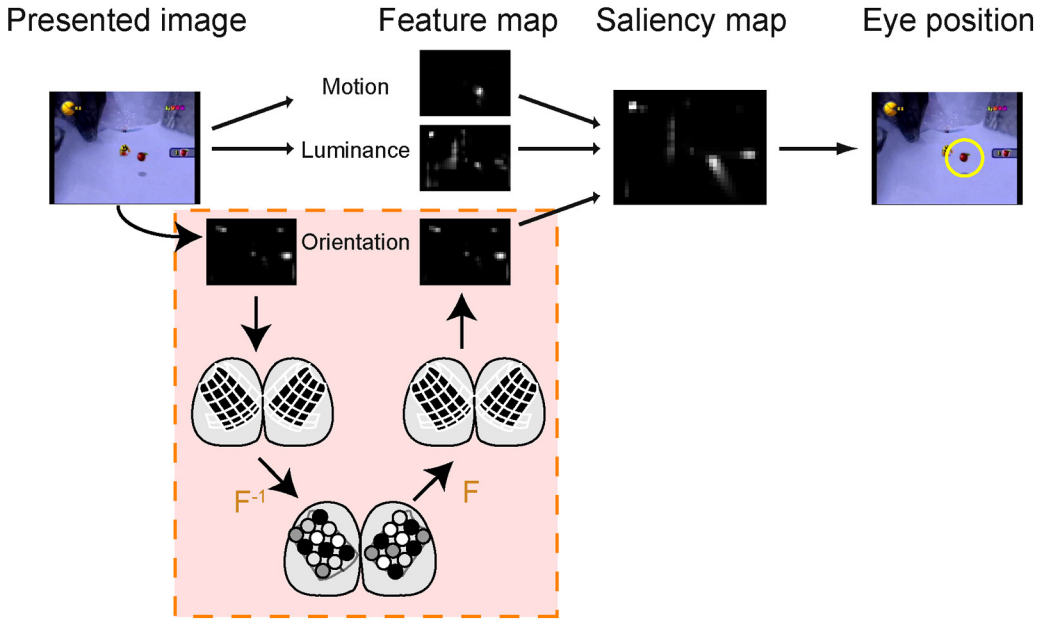

Fig. 4. Application in lesioned monkeys. The saliency computational model is able to predict the pattern of eye movements in monkeys during a free-viewing task. After V1 lesion, orientation saliency is abolished. The saliency-guided prosthetic system (orange) is used to restore orientation saliency by injecting currents into the SC to recover appropriate patterns of eye movements.

modeled as a saliency map that lacks an orientation feature map (Fig. 4, top) (Yoshida et al., 2012). Thus, the saliency-guided neural prosthesis presented in this paper (Fig. 4, orange box) can restore gaze function in monkeys with damage in V1 (Fig. 4), specifically by recovering the portions of gaze behavior caused by orientation saliency. In monkeys, we will be able to concretely evaluate the performance of the prosthesis by quantifying the contribution of orientation features to the subjects' gaze allocation during a free-viewing task such as the task used by Yoshida et al. (2012). The prosthesis is promising not only for subjects with visual deficits (retinal or cortical damage) but also for those with attentional deficits (hemineglect).

The objective of this paper was to describe the design of the prosthetic system and to demonstrate its feasibility via simulations. To demonstrate the operational principles of our system clearly, we used a model involving some major simplifying assumptions. These assumptions introduce limitations to the system that will need to be addressed in the future. We enumerate those limitations here. (1) Modeling of the response to electrical stimulation needs to be elaborated. The current system considers only the spatial pattern of evoked activity in SC but to accurately shape SC activity it will be necessary to incorporate temporal dynamics such as recurrent connectivity in the model predictions. (2) The current system assumes the spread of neural activity based on the previous research (Trappenberg et al., 2001), which is concerned with interaction between visual stimuli rather than electrical stimulation. Physical spread of electrical current (Tehovnik, 2006) must also be considered. (3) Electrodes can be placed precisely in simulation, but not in real application. (4) We do not know how the evoked SC activity reflecting orientation saliency may interact with other feature maps or other spontaneous SC activity (Fig. 4). We will address these problems in forthcoming simulations which build on the results presented in this paper using a spiking neural network model which explicitly models the dynamics of neurons in retina, V1, SC and other parts of the early visuomotor system (Veale, 2013).

Acknowledgements

M.Y. is funded by KAKENHI (22220006, 25430022) and Strategic Japanese–German Cooperative Programme from Japan Science and Technology Agency. R.V. is funded by an NSF Graduate Research Fellowship and NSF IGERT Traineeship.

References

Bradley, D.C., 2004. Visuotopic mapping through a multichannel stimulating implant in primate V1. Journal of Neurophysiology 93, 1659–1670.
Carello, C.D., Krauzlis, R.J., 2004. Manipulating intent: evidence for a causal role of the superior colliculus in target selection. Neuron 43, 575–583.
Chang, C.-K., Siagian, C., Itti, L., 2010. Mobile robot vision navigation & localization using gist and saliency. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2010, 4147–4154.
Dobelle, W.H., Mladejovsky, M.G., 1974. Phosphenes produced by electrical stimulation of human occipital cortex, and their application to the development of a prosthesis for the blind. The Journal of Physiology 243, 553–576.
Fecteau, J.H., Munoz, D.P., 2006. Salience, relevance, and firing: a priority map for target selection. Trends in Cognitive Sciences 10, 382–390.
Itti, L., 2005. Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. Visual Cognition 12, 1093–1123.
Itti, L., Koch, C., 2001. Computational modelling of visual attention. Nature Reviews Neuroscience 2, 194–203.
Moore, T., 2003. Microstimulation of the frontal eye field and its effects on covert spatial attention. Journal of Neurophysiology 91, 152–162.
Moore, T., Fallah, M., 2001. Control of eye movements and spatial attention. Proceedings of the National Academy of Sciences United States of America 98, 1273–1276.
Müller, J.R., Philiastides, M.G., Newsome, W.T., 2005. Microstimulation of the superior colliculus focuses attention without moving the eyes. Proceedings of the National Academy of Sciences United States of America 102, 524–529.
Navalpakkam, V., Itti, L., 2005. Modeling the influence of task on attention. Vision Research 45, 205–231.
Ottes, F.P., van Gisbergen, J.A., Eggermont, J.J., 1986. Visuomotor fields of the superior colliculus: a quantitative model. Vision Research 26, 857–873.
Schmidt, E.M., Bak, M.J., Hambrecht, F.T., Kufta, C.V., O'Rourke, D.K., Vallabhanath, P., 1996. Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex. Brain 119 (Pt. 2), 507–522.
Tehovnik, E.J., 2006. Direct and indirect activation of cortical neurons by electrical microstimulation. Journal of Neurophysiology 96, 512–521.
Thompson, K.G., Bichot, N.P., 2005. A visual salience map in the primate frontal eye field. Progress in Brain Research 147, 251–262.
Trappenberg, T.P., Dorris, M.C., Munoz, D.P., Klein, R.M., 2001. A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. Journal of Cognitive Neuroscience 13, 256–271.
Tseng, P.-H., Cameron, I.G.M., Pari, G., Reynolds, J.N., Munoz, D.P., Itti, L., 2012. High-throughput classification of clinical populations from natural viewing eye movements. Journal of Neurology, 1–10.
Veale, R., 2013. A neurorobotics approach to investigating word learning behaviors. In: Gogate, L., Hollich, G. (Eds.), Theoretical and Computational Models of Word Learning: Trends in Psychology and Artificial Intelligence. IGI Global, pp. 270–295.
Wallace, M.T., McHaffie, J.G., Stein, B.E., 1997. Visual response properties and visuotopic representation in the newborn monkey superior colliculus. Journal of Neurophysiology 78, 2732–2741.
Yoshida, M., Itti, L., Berg, D.J., Ikeda, T., Kato, R., Takaura, K., White, B.J., Munoz, D.P., Isa, T., 2012. Residual attention guidance in blindsight monkeys watching complex natural scenes. Current Biology 22, 1429–1434.